บทที่ 5

วิธีการศึกษา

5.1 การทดสอบคุณสมบัติความมีเสถียรภาพ (Stationary)

เนื่องจากข้อมูลที่ใช้ในการศึกษาครั้งนี้เป็นข้อมูลอนุกรมเวลา จึงต้องทำการตรวจสอบว่าข้อมูลมีลักษณะ stationary หรือไม่ เพราะตัวแปรเหล่านี้มักมีลักษณะ non-stationary หรือ stochastic process กล่าวคือ mean และ variance ของข้อมูลเหล่านั้นมีค่าเปลี่ยนแปลงไปตามระยะเวลา อาจก่อให้เกิดปัญหาทางด้าน spurious regression (Granger and Newbold, 1974) และ Phillipes (1988: อ้างถึงใน รังสรรค์ หทัยเสรี, 2538) สังเกตได้จากค่าสถิติบางอย่าง เช่น R^2 ที่ได้มีค่าสูงมากในขณะที่ D.W มีค่าต่ำมาก น่าจะเป็นเพราะ ตัวแปรมีความสัมพันธ์ต่อกันใน "ลักษณะของเงื่อนเวลา" (Correlated Trend) มากกว่าใน "ลักษณะพื้นฐานทางเศรษฐกิจ" (underlying economic relationship) มีข้อควรสังเกตว่า ในกรณีที่ข้อมูลอนุกรมเวลามีความสัมพันธ์กับเงื่อนเวลา (time trend) ค่าความเบี่ยงเบนโดยรวมที่คำนวณได้จากสมการถดถอย $\left[\sum_t (y_t - \bar{y})^2\right]$ จะคลาดเคลื่อนเพิ่มขึ้นตลอดเวลา เมื่อเทียบกับค่าเฉลี่ย y ด้วยเหตุนี้ค่า R^2 ที่คำนวณได้ซึ่งเท่ากับ $R^2 = 1 - \left[\sum_t \hat{e}_t^2 / \sum_t (y_t - \bar{y})^2\right]$ จึงสูงขึ้น และโน้มเข้าสู่ "หนึ่ง" เมื่อค่าในวงเล็บเข้าใกล้ศูนย์ ส่วนการที่ D.W. มีค่าที่ต่ำนั้นสะท้อนให้เห็นว่า "ตัวแปรความคลาดเคลื่อน" (error terms) มีความสัมพันธ์ซึ่งกันและกันอย่างมาก

การศึกษาครั้งนี้ใช้ unit root test ในการทดสอบ stationary โดยใช้การทดสอบ DF (Dicky-Fuller (DF) test) Dicky and Fuller (1981: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) และการทดสอบ ADF (Augmented Dicky-Fuller (ADF) test) (Said and Dickey 1984: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) สมมติฐานว่างของการทดสอบ DF (DF test) คือ $H_0: \rho = 1$ กำหนดสมการ

$$X_t = \rho X_{t-1} + \varepsilon_t \quad (5.1)$$

เรียกว่าการทดสอบ unit root ถ้า $|\rho| < 1$ ค่า X_t จะมีลักษณะ stationary และถ้า $\rho = 1$ แล้วค่า X_t จะมีลักษณะ nonstationary อย่างไรก็ตามการทดสอบนี้สามารถทำได้อีกทางหนึ่งซึ่งเหมือนกับสมการที่ (5.1) คือ

$$\Delta X_t = \theta X_{t-1} + \varepsilon_t \tag{5.2}$$

ซึ่งก็คือ $X_t = (1+\theta)X_{t-1} + \varepsilon_t$ ซึ่งก็คือสมการที่ (5.1) นั่นเอง โดยที่ $\rho = (1+\theta)$ ถ้า θ ในสมการที่ (5.2) มีค่าเป็นลบจะได้ว่า ρ ในสมการที่ (5.1) จะมีค่าน้อยกว่า 1 ดังนั้นสามารถจะสรุปได้ว่าการปฏิเสธ $H_0 : \theta = 0$ ซึ่งเป็นการยอมรับ $H_a : \theta < 0$ หมายความว่า $\rho < 1$ และ X_t มี integration of order zero (Charemza and Deadman, 1992) นั่นคือ X_t มีลักษณะstationary และถ้าเราไม่สามารถปฏิเสธ $H_0 : \theta = 0$ ได้ก็จะหมายความว่า X_t มีลักษณะ nonstationary

ถ้า X_t เป็น random walk with drift สามารถเขียนแบบจำลองได้ดังนี้

$$\Delta X_t = \alpha + \theta X_{t-1} + \varepsilon_t \tag{5.3}$$

และถ้า X_t มีแนวโน้มตามเวลาเชิงเส้น สามารถเขียนแบบจำลองได้ดังนี้

$$\Delta X_t = \alpha + \beta_t + \theta X_{t-1} + \varepsilon_t \tag{5.4}$$

โดยที่ $t =$ เวลา ซึ่งก็จะทำการทดสอบ $H_0 : \theta = 0$ โดยมี $H_a : \theta < 0$ เช่นเดียวกับที่กล่าวมาข้างต้น โดยสรุปแล้ว Dickey and Fuller (1979: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) ได้พิจารณาสมการถดถอย 3 รูปแบบที่แตกต่างในการทดสอบว่ามี unit root หรือไม่ ซึ่ง 3 สมการดังกล่าวได้แก่

$$\Delta X_t = \theta X_{t-1} + \varepsilon_t$$
$$\Delta X_t = \alpha + \theta X_{t-1} + \varepsilon_t$$
$$\Delta X_t = \alpha + \beta_t + \theta X_{t-1} + \varepsilon_t$$

โดยตัวพารามิเตอร์ที่อยู่ในความสนใจในทุกสมการคือ θ นั่นคือ ถ้า $\theta = 0$ แล้ว X_t จะมี unit root โดยการเปรียบเทียบค่าสถิติ $t(t-statistic)$ ที่คำนวณได้กับค่าที่เหมาะสมที่อยู่ในตาราง Dickey and Fuller (Enders, 1995) หรือกับค่าวิกฤติ MacKinnon Gujarati (1995: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546)

อย่างไรก็ตามค่าวิกฤติจะไม่เปลี่ยนแปลง ถ้าสมการ (5.2) (5.3) และ (5.4) ถูกแทนที่โดย autoregressive processes

$$\Delta X_t = \theta X_{t-1} + \sum_{i=1}^{\rho} \phi_i \Delta X_{t-1} + \varepsilon_t \tag{5.5}$$

$$\Delta X_t = \alpha + \theta X_{t-1} + \sum_{i=1}^{\rho} \phi_i \Delta X_{t-1} + \varepsilon_t \tag{5.6}$$

$$\Delta X_t = \alpha + \beta_t + \theta X_{t-1} + \sum_{i=1}^{\rho} \phi_i \Delta X_{t-1} + \varepsilon \tag{5.7}$$

(Enders, 1995 และ Gujarati, 1995) จำนวนของ lagged difference terms ที่จะนำเข้ามารวมในสมการนั้นจะมีมากพอที่จะทำให้พจน์ค่าความคลาดเคลื่อน มีลักษณะเป็น serially independent และเมื่อนำเอาการทดสอบ DF(Dickey-Fuller (DF) test) มาใช้กับสมการ (5.5)-(5.7) เราจะเรียกว่าการทดสอบ ADF (Augmented Dickey-Fuller (ADF) test) ค่าสถิติทดสอบ ADF มีการแจกแจงเชิงเส้นกำกับเหมือนกับสถิติ DF ดังนั้นก็สามารถใช้ค่าวิกฤติแบบเดียวกัน Gujarati (1995: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546)

5.2 การทดสอบความสัมพันธ์ในระยะยาว (Cointegration Test)

cointegration test เป็นการศึกษาเพื่อหาความสัมพันธ์ในระยะยาว (long – run relationship) ภายใต้เงื่อนไขของตัวแปรที่จะนำมาทดสอบ คือ ตัวแปรทุกตัวที่นำมาสร้างแบบจำลองนั้นต้องถูก integrate หรือมีคุณสมบัติ stationary ในอันดับเดียวกัน สำหรับในการศึกษานี้จะเลือกใช้วิธีของ Engle and Granger (1987: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546)

Engle และ Granger (1987: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) ได้ให้คำนิยามของ cointegration ของสองตัวแปรเป็นดังนี้คือ ถ้า x_t และ y_t เป็นอนุกรมเวลา (time series) x_t และ y_t จะถูกเรียกว่าเป็นอันดับของการร่วมกันไปด้วยกัน (cointegrated of order) d, b ซึ่งเขียนแทนด้วยสัญลักษณ์ $x_t, y_t \sim CI(d,b)$ ถ้า x_t และ y_t เป็น integrated of order d ซึ่งเขียนแทนด้วยสัญลักษณ์ $I(d)$ และจะต้องมีการรวมเชิงเส้น (linear combination) ของตัวแปรทั้งสองนี้ สมมุติว่าเป็น $\alpha x_t + \beta y_t$ ซึ่งจะต้องเป็น integrated of order $(d-b)$ โดยที่ $d > b > 0$ เวกเตอร์ $[\alpha, \beta]$ นี้จะถูกเรียกว่า เวกเตอร์ที่ทำให้เกิดการร่วมกันไปด้วยกัน (cointegrating vector) Charemza and Deadman (1992: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) ยกตัวอย่างเช่น ถ้า x_t และ y_t เป็น $I(1)$

ทั้งคู่ และพจน์ค่าความคลาดเคลื่อน (error term) ε_t ของการถดถอยเชิงเส้น (linear regression) ของตัวแปรทั้งสองเป็นกระบวนนิ่ง (stationary process) $I(0), x_t$ และ y_t จะถูกเรียกว่าเป็นอันดับของการร่วมกันไปด้วยกัน (cointegrated of order) (1,1) หรือ $x_t, y_t \sim CI(1,1)$ เพราะฉะนั้นการถดถอยร่วมกันไปด้วยกัน (cointegration regression) ก็คือ เทคนิคการประมาณค่าความสัมพันธ์ดุลยภาพระยะยาว (long–term equilibrium relationship) ระหว่างอนุกรมที่มีลักษณะไม่นิ่ง (nonstationary series) โดยการเบี่ยงเบน (deviations) จากวิถีดุลยภาพระยะยาว (long–term equilibrium path) นี้มีลักษณะนิ่ง (stationary) (Ling *et al.*, 1998)

อย่างไรก็ตาม ถ้า x_t คือ เวกเตอร์ n×1 (n×1 vector) ของอนุกรม $x_{1t}, x_{2t}, \ldots, x_{nt}$ และถ้าแต่ละ x_t เป็น $I(d)$ โดยที่ $i = 1,\ldots,n$ และมี α ซึ่งคือ เวกเตอร์ n×1 (n×1 vector) ที่ทำให้ $X_t'\alpha \sim I(d-b)$ ดังนั้น $X_t'\alpha \sim CI(d-b)$

สำหรับในทางเศรษฐมิติเชิงประจักษ์แล้วกรณีที่น่าสนใจที่สุด คือ กรณีที่อนุกรม (series) ที่ถูกแปลง (transformed) ด้วยเวกเตอร์ที่ทำให้เกิดการร่วมกันไปด้วยกัน (cointegrating vector) มีลักษณะนิ่ง (stationary) นั่นคือ กรณีที่ $d = b$ และสัมประสิทธิ์ของการร่วมกันไปด้วยกัน (cointegrating coefficients) สามารถที่จะหาออกมาได้ด้วยพารามิเตอร์ที่อยู่ในสมการความสัมพันธ์ระยะยาวระหว่างตัวแปรต่างๆ ในแบบจำลอง Charemza and Deadman (1992: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546)

สำหรับการทดสอบการร่วมกันไปด้วยกัน (cointegration) นั้น ให้ใช้ส่วนตกค้างหรือส่วนที่เหลือ (residuals) จากสมการถดถอย (regression equation) ที่เราต้องการทดสอบการร่วมกันไปด้วยกัน (cointegration) ซึ่งคือ $\hat{e}_t$ มาทำการถดถอยดังสมการดังต่อไปนี้

$$\Delta\hat{e}_t = \gamma\hat{e}_{t-1} + v_t \tag{5.8}$$

Gujarati (1995: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) และนำค่าสถิติ t (t–statistic) ซึ่งได้มาจากอัตราส่วนของ $\hat{\gamma} / \text{S.E.}\,\hat{\gamma}$ ไปเปรียบเทียบกับค่าวิกฤติ MacKinnon (MacKinnon critical values) โดยที่สมมุติฐานว่างของการไม่มีการร่วมกันไปด้วยกัน (null hypothesis of no cointegration) คือ H_0 : $\gamma = 0$ ค่าลบของค่าสถิติ t (t–statistic) ที่มีนัยสำคัญก็จะเป็นการปฏิเสธ H_0 ซึ่งก็จะนำไปสู่ข้อสรุปว่าตัวแปรที่มีลักษณะไม่นิ่ง (nonstationary) ในสมการดังกล่าวร่วมกันไปด้วยกัน (cointegrated) (Johnston and Dinardo ,1997)

อย่างไรก็ตาม ถ้าส่วนตกค้างหรือส่วนที่เหลือ (residuals) ของสมการ (5.8) ไม่เป็น white noise เราก็จะใช้การทดสอบ ADF (Augmented Dickey–Fuller (ADF) test) แทนที่จะใช้สม

การ (5.8) สมมุติว่า v_t ของสมการที่ (5.8) มีสหสัมพันธ์เชิงอันดับ (serial correlation) เราก็จะใช้สมการดังนี้

$$\Delta\hat{e}_t = \gamma\hat{e}_{t-1} + \sum_{i=1}^{p} a_i \Delta\hat{e}_{t-1} + v_t \tag{5.9}$$

และถ้า $-2 < \gamma < 0$ เราสามารถจะสรุปได้ว่า ส่วนตกค้างหรือส่วนที่เหลือ (residuals) มีลักษณะนิ่ง (stationary) และ x_t และ y_t จะเป็น $CI(1,1)$ โปรดสังเกตว่าสมการ (5.8) และ (5.9) ไม่มีพจน์ส่วนตัด (intercept term) เนื่องจาก $\hat{e}_t$ เป็นส่วนตกค้างหรือส่วนที่เหลือ (residuals) จากสมการถดถอย (regression equation) (Enders, 1995)

5.3 การทดสอบการปรับตัวในระยะสั้น (Error Correction Model)

ถ้า x_t และ y_t ร่วมกันไปด้วยกัน (cointegrated) ก็หมายความว่า ตัวแปรทั้งสองมีความสัมพันธ์เชิงดุลยภาพระยะยาว (long term equilibrium relationship) แต่ในระยะสั้นอาจจะมีการออกนอกดุลยภาพ (disequilibrium) ได้ เพราะฉะนั้นเราสามารถจะให้พจน์ค่าความคลาดเคลื่อน (error term) ในสมการที่ร่วมกันไปด้วยกัน (cointegrated) เป็นค่าความคลาดเคลื่อนดุลยภาพ (equilibrium error) และเราสามารถที่จะนำเอาพจน์ค่าความคลาดเคลื่อน (error term) นี้ไปผูกพฤติกรรมระยะสั้นกับระยะยาวได้ Gujarati (1995: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) ลักษณะสำคัญของตัวแปรร่วมกันไปด้วยกัน (cointegrated variables) ก็คือว่าวิถีเวลา (time path) ของตัวแปรเหล่านี้จะได้รับอิทธิพลจากการเบี่ยงเบน (deviations) จากดุลยภาพระยะยาว (long–run equilibrium) และถ้าระบบจะกลับไปสู่ดุลยภาพระยะยาว (long–run equilibrium) การเคลื่อนไหวของ ตัวแปรอย่างน้อยบางตัวแปรจะต้องตอบสนองต่อขนาดของการออกนอกดุลยภาพ (disequilibrium) ใน error correction model (ใช้ชื่อย่อเช่นเดียวกันว่า ECM ซึ่งขึ้นอยู่กับความหมายในตอนนั้นว่าจะเน้นตรง mechanism หรือ model แต่ก็จะมีแนวคิดที่ใกล้เคียงกันมาก ตำราบางเล่มเรียก error correction model (ECM) บางเล่มเรียก error correction mechanism (ECM)) พลวัตพจน์ ระยะสั้น (short – term dynamics) ของตัวแปรในระบบจะได้รับอิทธิพลจากการเบี่ยงเบน (deviation) จาก ดุลยภาพ สำหรับแบบจำลอง ECM ที่เสนอโดย Ling *et al.* (1998 : อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) สามารถเขียนได้ดังนี้

$$\Delta y_t = a_1 + a_2\hat{e}_{t-1} + a_3\Delta x_t + \sum_{h=1}^{p} a_{4h}\Delta x_{t-h} + \sum_{l}^{q} a_{5l}\Delta y_{t-l} + \mu_t \tag{5.10}$$

โดยที่ $\hat{e}_t$ คือ ส่วนตกค้างและส่วนที่เหลือ (residuals) ของสมการการถดถอยร่วมกันไปด้วยกัน (cointegrating regression equation) ค่า a_2 จะให้ความหมายว่า a_2 ของความคลาดเคลื่อน (discrepancy)

ระหว่างค่าสังเกตที่เกิดขึ้นจริง (actual) ของ y_t กับค่าที่เป็นระยะยาว (long run) หรือดุลยภาพ (equilibrium) ในคาบ (period) ที่แล้วจะถูกขจัดไป (eliminated) หรือถูกแก้ไขไป (corrected) ในแต่ละคาบ (period) ต่อมา Gujarati (1995: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) เช่น ในแต่ละเดือน แต่ละสัปดาห์ หรือแต่ละไตรมาส นั่นคือ a_2 คือ สัดส่วนของการออกของดุลยภาพ (disequilibrium) ของ y ในคาบ (period) นี้ที่ถูกขจัดไปในคาบ (period) ต่อไป เป็นต้น

สำหรับรูปแบบ ECM ที่อ้างโดย Gujarati (1995: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) นั้น สามารถเขียนได้ ดังนี้

$$\Delta y_t = a_1 + a_2 \hat{e}_{t-1} + a_3 \Delta x_t + \mu_t \qquad (5.11)$$

แต่รูปแบบ ECM ที่กล่าวถึงโดย Charemza and Deadman (1992: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) ไม่มีพจน์คงที่ (constant term) และล้าหลัง (lagged) Δx ซึ่งสามารถแสดงได้ดังนี้

$$\Delta y_t = a_1 \hat{e}_{t-1} + a_2 \Delta x_t + \mu_t \qquad (5.12)$$

โดยที่ a_1 มีค่าเป็นลบ

อย่างไรก็ตาม Enders (1995: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) ระบุ error correction model (ECM) ดังนี้

$$\Delta y_t = a_1 + a_2 \hat{e}_{t-1} + \sum_{h=1}^{p} a_{4h} \Delta x_{t-h} + \sum_{l=1}^{q} a_{5l} \Delta y_{t-l} + \mu_{yt} \qquad (5.13)$$

$$\Delta x_t = b_1 + b_2 \hat{e}_{t-1} + \sum_{m=1}^{r} b_{4m} \Delta x_{t-m} + \sum_{n=1}^{s} b_{5n} \Delta y_{t-n} + \mu_{xt} \qquad (5.14)$$

โดยที่ไม่มีตัวแปร Δx_t ในสมการที่ (5.13) และ Δy_t ในสมการที่ (5.14) ซึ่งแตกต่างไปจากแบบจำลองที่ใช้โดย Ling *et al.* (1998)

Tambi (1999: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) ได้สร้าง error correction model โดยมีสมการเดียวและภายในสมการดังกล่าวจะเหมือนกันกับ สมการที่ (5.13)

5.4 การประมาณค่าด้วยวิธี The Generalized Method of Moments

ผู้ที่เสนอวิธีการแบบ Generalized Method of Moments (GMM) คือ Hansen (1982: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) วิธีการนี้เป็นการประมาณค่าพารามิเตอร์ของแบบจำลองโดยตรงจากเงื่อนไขโมเมนต์ (moment conditions) ซึ่งใส่เข้ามาในแบบจำลอง เงื่อนไขเหล่านี้สามารถที่จะมีลักษณะเชิงเส้น (linear) ในพารามิเตอร์ แต่บ่อยครั้งมากทีเดียวจะมีลักษณะไม่เชิงเส้น (nonlinear) และเพื่อที่จะทำให้เราสามารถหาค่าพารามิเตอร์ได้ จำนวนของเงื่อนไขโมเมนต์อย่างน้อยที่สุดควรจะเท่ากับจำนวนพารามิเตอร์ที่ไม่ทราบค่า (unknown parameters) (Verbeek, 2000)

ในกรณีที่เงื่อนไขโมเมนต์ไม่เป็นเชิงเส้น คือมีลักษณะเป็น nonlinear ตัวประมาณค่าตัวแปรเครื่องมือหรือที่เรียกว่า instrumental variables estimator ก็จะมีปัญหา Verbeek (2000: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) ได้ยกตัวอย่างว่าทำไมทฤษฎีทางเศรษฐศาสตร์จึงได้ให้นัย (imply) ถึงเงื่อนไขโมเมนต์ที่ไม่เชิงเส้น (nonlinear moment conditions) ตัวอย่างที่ Verbeek (2000: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) ยกมานั้นได้มาจากการศึกษาของ Hansen และ Singleton (1982: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) ซึ่งสมมุติให้ปัจเจกบุคคลต้องการทำให้ความพอใจที่คาดหมาย (expected utility) ของการบริโภคในปัจจุบันและในอนาคตมีค่าสูงสุด โดยหาค่าจากสมการดังต่อไปนี้

$$Max\, E_t \left\{ \sum_{s=0}^{s} \delta^s U\left(C_{t+s}\right) \right\} \tag{5.15}$$

โดยที่ C_{t+s} = การบริโภคในคาบที่ t + s

$U(C_{t+s})$ = ความพอใจซึ่งขึ้นอยู่กับระดับของการบริโภคซึ่งถูก discounted โดย discount factor $\delta \left(0 < \delta \leq 1\right)$

E_t = expectation operator ซึ่งขึ้นอยู่กับข้อมูลทั้งหมดที่มีอยู่ ณ เวลา t และในกรณีนี้ข้อจำกัดทางด้านงบประมาณก็จะสามารถเขียนได้ดังนี้

$$C_{t+s} + q_{t+s} = W_{t+s} + \left(1 + r_{t+s}\right) q_{t+s-1} \tag{5.16}$$

โดยที่ q_{t+s} = ความมั่งคั่งทางการเงิน (financial wealth) ณ สิ้นสุดของคาบ t + s

r_{t+s} = ผลตอบแทนของความมั่งคั่งทางการเงิน (ที่ได้ลงทุนใน portfolio ของทรัพย์สิน)

W_{t+s} = รายได้จากแรงงาน (labor income)

ข้อจำกัดทางด้านงบประมาณนี้กล่าวว่า รายได้จากแรงงานบวกด้วยรายได้จากทรัพย์สินควรจะถูกนำไปใช้จ่ายหรือซื้อเพื่อการบริโภคเท่ากับ C_{t+s} หรือออมไว้เท่ากับ q_{t+s} ปัญหาการทำให้มีค่าสูงสุดปัญหานี้ยากมากที่จะหาค่าได้ อย่างไรก็ตามยังคงเป็นไปได้ที่จะประมาณค่าพารามิเตอร์ที่ไม่ทราบค่าที่เกี่ยวข้องโดยผ่านเงื่อนไขอันดับที่หนึ่ง (first order conditions) เงื่อนไขอันดับหนึ่งของสมการที่ (5.15) ภายใต้สมการที่ (5.16) ให้นัยว่า

$$E_t\left\{\delta U'(C_{t+1})(1+r_{t+1})\right\} = U'(C_t) \tag{5.17}$$

โดยที่ U' = อนุพันธ์ที่หนึ่งของ U

ทางขวามือของสมการที่ (5.17) ก็คือ ความพอใจส่วนเพิ่ม (marginal utility) ของหนึ่งบาทที่เพิ่มขึ้นที่ใช้ในการบริโภคในวันนี้ ในขณะนี้ทางซ้ายมือก็จะให้ความพอใจส่วนเพิ่มที่คาดหมาย (expected marginal utility) ของการออมเงินหนึ่งบาทนี้จนกระทั่งคาบต่อไป (ซึ่งจะกลายเป็น $1+r_{t+1}$ บาท) และจึงจะบริโภค จุดที่เหมาะสมก็คือ จุดที่แสดงความเท่ากันระหว่างความพอใจส่วนเพิ่ม (marginal utilities) (ที่คาดหมาย (expected))

เราสามารถจะเขียนสมการที่ (5.17) ใหม่ได้ดังนี้

$$E_t\left\{\frac{\delta U'(C_{t+1})}{U'(C_t)}(1+r_{t+1})-1\right\} = 0 \tag{5.18}$$

นี่คือ เงื่อนไขโมเมนต์ (moment condition) หนึ่งเงื่อนไขที่เราสามารถที่จะนำไปใช้ประมาณค่าตัวพารามิเตอร์ที่ไม่ทราบค่าได้ ถ้าเรามีข้อสมมุติเกี่ยวกับฟังก์ชันความพอใจ (utility function) U เราสามารถทำได้ดังนี้ ในการประมาณค่าพารามิเตอร์ที่ไม่ทราบค่า สมมุติให้ z_i ได้ถูกรวมอยู่ในเซตของข้อมูล ในกรณีนี้จะให้ความหมายว่า z_i ไม่ได้ให้ข้อมูลใดๆ เกี่ยวกับค่าคาดหมาย (expected value) ของ

$$\frac{\delta U'(C_{t+1})}{U'(C_t)}(1+r_{t+1})-1$$

เพราะฉะนั้นเราจะได้ว่า

$$E\left\{\left(\frac{\delta U'(C_{t+1})}{U'(C_t)}(1+r_{t+1})-1\right)z_t\right\} = 0 \qquad (5.19)$$

ในกรณีนี้เราสามารถจะกล่าวได้ว่า z_i คือ เวกเตอร์ (vector) ของเครื่องมือ (instruments) ซึ่งถูกต้องโดยข้อสมมุติเกี่ยวกับพฤติกรรมที่เหมาะสมหรือดีที่สุด (optimal behavior) (การคาดหมายอย่างมีเหตุผล) ของปัจเจกบุคคลนั้น

เพื่อให้เกิดความง่าย เราจะสมมุติว่าฟังก์ชันความพอใจมีลักษณะดังนี้

$$U(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma} \qquad (5.20)$$

โดยที่ γ = สัมประสิทธิ์ (ค่าคงตัวหรือค่าคงที่) แห่งการหลีกเลี่ยงความเสี่ยงสัมพัทธ์ (relative risk aversion) โดยที่ค่าที่สูงขึ้นของ γ นั้นหมายความว่า ปัจเจกบุคคลนั้นจะหลีกเลี่ยงความเสี่ยงมากขึ้น

ถ้าฟังก์ชันความพอใจมีลักษณะสมการที่ (5.20) เราก็สามารถเขียนสมการที่ (5.19) ได้ดังนี้

$$E\left\{\left(\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}(1+r_{t+1})-1\right)z_t\right\} = 0 \qquad (5.21)$$

ในตอนนี้เราก็จะมีเซตของเงื่อนไขโมเมนต์ (moment conditions) ซึ่งทำให้เราสามารถหาค่าตัวพารามิเตอร์ที่ไม่ทราบค่าคือ δ และ γ ได้ โดยกำหนดค่าสังเกต (observations) เกี่ยวกับ c_{t+1}/c_t , r_{t+1} และ z_t มาให้ ทำให้เราสามารถประมาณค่าพารามิเตอร์ได้อย่างคล้องจอง (consistently) (Verbeek, 2000)

ต่อไปจะพิจารณาแบบจำลองซึ่งมีเซต (set) ของ R เงื่อนไขโมเมนต์ (moment conditions) ดังนี้

$$E\{f(W_t, Z_t, \theta)\} = 0 \tag{5.22}$$

โดยที่ f = ฟังก์ชันเวกเตอร์ (vector function) ซึ่งประกอบไปด้วย R สมาชิก (elements)

θ = เวกเตอร์ (vector) ที่มีมิติ (dimension) เท่ากับ K ซึ่งเป็นเวกเตอร์ (vector) ของพารามิเตอร์ที่ไม่ทราบค่า

W_t = เวกเตอร์ (vector) ของตัวแปรที่สามารถสังเกตได้ซึ่งอาจจะเป็นตัวแปรในระบบ (endogenous) หรือตัวแปรนอกระบบ (exogenous) ก็ได้

Z_t = เวกเตอร์ (vector) ของเครื่องมือ (instruments)

และตามตัวอย่างข้างต้น $w_t' = (c_{t+1}/c_t, r_{t+1})$ เป็นต้น

ในการประมาณค่า θ เราจะใช้วิธีการแบบเดียวกับวิธีการข้างต้น และเราจะใช้ตัวอย่างสมมูล (equivalent) กับสมการที่ (5.22) ซึ่งกำหนดโดย

$$g_T(\theta) \equiv \frac{1}{T}\sum_{t=1}^{T} f(W_t, Z_t, \theta) \tag{5.23}$$

ถ้าจำนวนของเงื่อนไขโมเมนต์ (moment conditions) ซึ่งมีเท่ากับ R เงื่อนไขเท่ากันกับจำนวนของพารามิเตอร์ที่ไม่ทราบค่า K ตัว เราก็จะสามารถให้ R สมาชิก (elements) ในสมการ (5.9) เท่ากับศูนย์ และหาค่า θ ออกมาซึ่งจะได้ตัวประมาณค่าที่คล้องจอง (consistent estimator) ที่มีลักษณะหนึ่งเดียวหรือเป็นไปได้อย่างเดียว (unique) (Verbeek, 2000) นอกจากนี้ Verbeek (2000: อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2546) ยังได้ให้ข้อสังเกตเพิ่มเติมดังนี้

ถ้า f มีลักษณะไม่เชิงเส้น (nonlinear) ใน θ ผลเฉลย (solution) อาจจะไม่สามารถหาได้ และถ้าจำนวนเงื่อนไขโมเมนต์ (moment conditions) น้อยกว่าจำนวนพารามิเตอร์เราจะไม่สามารถหาค่าพารามิเตอร์ที่ไม่ทราบค่าแบบหนึ่งเดียว (uniquely) ได้โดยการให้สมการที่ (5.23) มีค่าเท่ากับศูนย์

ทำการเลือกตัวประมาณค่า (estimator) สำหรับ θ ในลักษณะที่ว่า เวกเตอร์ (vector) ของโมเมนต์ตัวอย่าง (sample moments) มีค่าใกล้ศูนย์เท่าที่จะเป็นได้ในความหมายที่ว่ารูปแบบกำลังสอง (quadratic form) ใน $g_T(\theta)$ มีค่าต่ำสุด นั่นคือ

$$\min_{\theta} Q_T(\theta) \;=\; \min_{\theta} g_T(\theta)' W_T\, g_T(\theta) \qquad (5.24)$$

โดยที่ W_T = เมทริกซ์บวกแน่นอน (positive definite matrix ด้วย $p\lim W_T = W$ ผลเฉลย (solution) ของปัญหานี้คือ generalized method of moments หรือ GMM estimator θ และเราก็สามารถแสดงให้เห็นได้ว่าตัวประมาณค่า GMM (GMM estimator) มีลักษณะคล้องจอง (consistent) และมีการแจกแจงปกติเชิงเส้นกำกับภายใต้ weak regularity conditions

ในทางปฏิบัติตัวประมาณค่า GMM (GMM estimator) นั้นหาได้จากการหาค่าจากการ minimize สมการที่ (5.24) โดยวิธีการต่างๆ ดังมีรายละเอียดใน Greene (2000) Verbeek (2000)

ดังทราบกันดีอยู่แล้วว่า เมทริกซ์ที่ใช้ถ่วงน้ำหนัก W_T ซึ่งแตกต่างกัน จะให้ตัวประมาณค่าคล้องจอง (consistent estimators) ที่แตกต่างกัน เมทริกซ์ที่ใช้ถ่วงน้ำหนักที่เหมาะสมซึ่งจะนำไปสู่เมทริกซ์ความแปรปรวนของโมเมนต์ตัวอย่าง (sample moments) ในกรณีที่ไม่มี autocorrelation เมทริกซ์ที่ถ่วงน้ำหนักที่เหมาะสมสามารถเขียนได้ดังนี้

$$W^{opt} \;=\; \left(E\left\{f(W_t, Z_t, \theta) f(W_t, Z_t, \theta)'\right\}\right)^{-1}$$

โดยทั่วไปแล้วเมทริกซ์นี้จะขึ้นอยู่กับเวกเตอร์ของพารามิเตอร์ที่ไม่ทราบค่า θ ซึ่งเป็นปัญหาที่เราไม่ได้พบในแบบจำลองเชิงเส้น ผลเฉลย (solution) ก็คือ เราจะใช้กระบวนการประมาณค่าหลายขั้นตอน ในขั้นตอนแรกเราจะใช้ suboptimal choice ของ W_T ซึ่งไม่ขึ้นอยู่กับ θ เช่น เมทริกซ์เอกลักษณ์ (identity matrix) เพื่อที่จะหาค่าตัวประมาณค่าคล้องจอง (consistent estimator) ตัวแรก $\hat{\theta}_{[I]}$ หลังจากนั้นเราก็จะประมาณค่าเมทริกซ์ที่ใช้ถ่วงน้ำหนักที่เหมาะสมโดย

$$W^{opt} \;=\; \left(\frac{1}{T}\sum_{t=1}^{T} f\left(W_t, Z_t, \hat{\theta}_{[1]}\right) f\left(W_t, Z_t, \hat{\theta}_{[1]}\right)'\right)^{-1} \qquad (5.25)$$

ในขั้นตอนที่สองเราก็จะสามารถหาตัวประมาณค่า GMM ที่มีประสิทธิภาพ (เหมาะสม) เชิงเส้นกำกับ $\hat{\theta}_{GMM}$ ได้ โดยมีการแจกแจงเชิงเส้นกำกับ (asymptotic distribution) ดังนี้

$$\sqrt{T}\left(\hat{\theta}_{GMM} - \theta\right) \;\to\; N(0, v) \qquad (5.26)$$

โดยที่เมทริกซ์ความแปรปรวนร่วมเชิงเส้นกำกับ (asymptotic covariance matrix) v คือ

$$v = \left(DW^{opt}D'\right)^{-1} \tag{5.27}$$

โดยที่ D คือ k x R เมทริกซ์อนุพันธ์ (derivative matrix)

$$D = E\left\{\frac{\partial f(W_t, Z_t, \theta)}{\partial \theta'}\right\} \tag{5.28}$$

สำหรับ over identifying restrictions test สำหรับแบบจำลองไม่เชิงเส้น (nonlinear models) นั้น ถ้ากำหนดให้ว่า เงื่อนไขโมเมนต์ (moment conditions) นั้นถูกต้อง สถิติทดสอบ (test statistic)

$$\xi = Tg_T\left(\hat{\theta}_{GMM}\right)' W_T^{opt} g_T\left(\hat{\theta}_{GMM}\right)$$

โดย $\hat{\theta}_{GMM}$ คือ ตัวประมาณค่า GMM ที่เหมาะสม และ W_T^{OPT} คือ เมทริกซ์ที่ใช้ถ่วงน้ำหนักที่เหมาะสมในสมการที่ (5.25) ζ นี้มีการแจกแจงเชิงเส้นกำกับแบบไคสแควร์ด้วย $R-K$ degrees of freedom ถ้าในกรณีที่เป็น exactly identified ก็จะมี degrees of freedom เป็นศูนย์ เราก็จะไม่มีอะไรจะต้องทดสอบ (Verbeek, 2000)